حول تأليل المقاطع التركيبية الجامدة في اللغات الطبيعية

إعداد محمد عمر أمطوش

أستاذ ملحق بالمجلس العلمي بمراكش

أكيدا لن يتفق الكثيرون على فحوى النص التالي، بل قد يوجد من سيرفضه أساسا، بدعوى ما الفائدة من نقل علوم وأفكار ليست من تراثنا ولا فائدة منها خاصة وأن الموضوع يتناول اللغة وحيث أن جل ما يأتينا من الغرب في هذا الميدان هو شديد الارتباط بحالة اللغات الغربية فلا ينتظر أن تقدم شيئا جديدا للعربية، اللغة التي تختلف عنها. وبالطبع بعض المنتقدين أغلق الله بصره وبصيرته على كل ما هو جديد وبعضهم جهلا بمرامي المجال ودقته يتيه في جزئيات ثانوية.

تدفع المعالجة الآلية للغات بالعاملين في المجال إلى ولوج أبواب متميزة قد تكون مجهولة أو ربما متقادمة عند أصحاب اللسانيات ومن تلك الأبواب ما سمي "بالصيغ الجامدة[1]" ونسميه هنا "بالوحدات المعجمية المركبة": وفعلا فقد لوحظ في سنوات الخمسينات، مع بداية تأليل اللغات، عودة لصدارة الاهتمامات لما نسميه الوحدات المعجمية المركبة وقد تعددت عند النقلة العرب أو غيرهم مسميات هذه الوحدات فمن يستعمل للدلالة على مفهوم الوحدات المعجمية المركبة

---

[1] قد يجد غير المهتمين بمجالات اللسانيات أو المعالجات الآلية صعوبة في تتبع مضامين المصطلحات المستعملة هنا وسنحاول قدر الإمكان التنبيه لذلك في الهوامش أو بتغيير الخط في صلب النص، علما بأننا لم نختر دوما ذكر المصطلح الغربي مع مقابله الأجنبي تأصيلا للموضوع عند العرب، زيادة على أن الساحة العربية تزخر بمعاجم تسمح بمعرفة المقابل.

مصطلحات الجمل أو الصيغ المسبوكة والتعبيرات أو التراكيب الجامدة وسمات جملية ومقاطع جامدة أو مسبوكة وغيرها؛ وفي الأدبيات الأمريكية طغت كلمة "أيديومز"[1] بتأثير وتعميم من "بار هيليل"[2] سنة 1955؛ وهي استعمالات تعبر، أحيانا، عن اجتهادات شخصية لفظية بحتة أو دلالة على إحدى السمات اللسانية الخاصة بهذه المقاطع أحيانا أخرى؛ وتجدر الإشارة إلى أن هذا الاهتمام اتخذ أشكالا وزخما متغيران حسب التقاليد والمدارس اللسانية؛ فهناك تباين في الاهتمام بهذه الصيغ شكلا وحيزا، ما بين المدارس الفرنسية والروسية والأمريكية والانجليزية، وبالطبع . وللأسف فالمدرسة العربية ليست إلا تابعا لواحدة منها في هذا المجال. ومن زاوية أخرى يجب التمييز بين توجهين في المعالجة الآلية للوحدات المعجمية المركبة توجه يستند، حسب ما يذكر هابير وجاكمان[3]، على منطلقين وهما "التحليل البنيوي" و"التحليل الإحصائي" وفي الواقع هذا التقسيم هو موضوع نقاش ومحط جدل وربما ألزمنا ذاك الجدل بالعودة للمنابع الفكرية والفرضيات والمناهج لمعرفة السبب. والسؤال البديهي الذي يفرض نفسه هو: لماذا برز الاهتمام بالوحدات المعجمية المركبة مع المعالجة الآلية للغات وارتباطا بالتساؤل السابق: هل كل معالجة آلية للغات تتطلب معالجة للوحدات المعجمية المركبة؟ ولتوضيح الفكرة يمكن أن نعطي بداية فكرة عما يميز المنطلق الإحصائي عن نظيره البنيوي مع معاينة تجارب معالجة الوحدات المعجمية المركبة في بدايات التأليل أو الأعمال التي نتجت

---

[1] Idioms

[2] Bar-Hillel יהושע בר-הלל ولد في فيينا عام 1915 ومات بالقدس فيلسوف ولساني ورياضي من أتباع ردولف كارناب في المنطق الرياضي، وتأثر بالخصوص بنظريته حول منطق التركيب اللغوي. وكان لبار هيليل اهتمام كبير بما يتعلق بتأليل اللغات فقد كان الباحث الأول الذي وظف في معهد مساشوسيت الأمريكي مكلفا حصرا بالترجمة الآلية وأنهى حياته في الجامعة العربية.

[3] Habet & Jacquemin

مباشرة عن هذه المعالجة : كتحليل *الترابط الانتظامي* عند مدرسة لندن أو الدراسات الإحصائية للغة الفرنسية، وهي أبحاث شديدة الارتباط بالمدرسة البنيوية الأمريكية والأوروبية، وكذا الأبحاث الروسية على *التمثيل الدلالي* في المعالجات الآلية. وهي مقاربات تعبر عن مواقف متباينة بخصوص المعجمة ومستوى حركيتها .

فنحن أمام مقاربتين للمجال الأولى ترى أولوية اعتبار الوحدات المعجمية المركبة قبل كل معالجة آلية عكس الثانية التي تعتبر معالجة الوحدات المعجمية المركبة ثانوية إذ لم نقل غير معنوية، فاعتبار منزلة مجمعات (تكتلات) الكلمات بالنسبة لها ترف. وفي أولي المقاربتين، وهي نابعة من البنيوية، نميز بين العديد من المواقف ولكنها تشترك في كثير من النقاط :

1) ضرورة تحديد موقع الوحدات المعجمية المركبة مقارنة *بالمركبات الحرة*[1] وهذا التحديد هو ركن أساسي لكل معالجة آلية؛

2) ضرورة ترتيب الوحدات المعجمية المركبة؛

3) تدوين الوحدات المعجمية المركبة في لوائح (معاجم) بغية استغلالها؛

4) التفريق بين من جانب مهمة تحديد وترتيب وتدوين الوحدات المعجمية المركبة ومن جانب آخر مهمة الاستغلال كالترجمة الآلية والبحث عن المعلومات في النص وغير ذلك؛ وهو ما استلزم في بعض الأبحاث المتأخرة معالجة أولية للنص.

ونقاط التباين في حيز هذه المقاربة يتلخص في:

---

[1] نفترض مقطعا ما (س ـ ص) إذا كان تعويض عنصر من عناصر المقطع بعنصر مشابه ممكنا ومقبولا لغة واستعمالا يكون المقطع حرا، أما إذا كان التكوين مضيقا أو مرفوضا فلا يكون المقطع وتركيبه حرا. وكون المقطع حرا أو لا له تأثير على مسار المعالجة ومقاربتها.

1) تباين الأهمية الموجهة للتحليل الداخلي لمجمعات الكلمات أي البحث عن قواعد تكونها وتركيبها؛

2) تفضيل التحليل *الصرفي* أو *التركيبي* في قواعد التكوين مع العلم بأن الأفضلية عموما أوليت للتركيب. ويستحسن معاينة الرجوع هنا للأعمال السباقة عند أرسين دارمستيتر[1] حول نمط تكوين التراكيب ؛

3) نمط تحديد المجمعات المدروسة كمعطيات لغة مجدولة مسبقا أو محددة داخل سلسلة مكتوبة بطرق تعرف آلي. وحسب الحالة فقد ننطلق من وحدات نعتبرها معجَّمة وهي مدونة في *لوائح مغلقة*[2] ويعني أن بقية السلاسل مركبات ومقاطع حرة ضرورة أو أن نعتبر بأن الوحدة المركبة هي محددة بالنظر لمتن *ولخطاب* معين.

4) منح أهمية أو لا لمجمعات الكلمات كوحدات معنوية معبرة، ففي الترجمة الآلية وعموما في مسار *توليد النص* يتوجب إيجاد مقابل في اللغة الهدف وغالبا ما يتوجب إعادة تشكيل المركب، ومشكل عدم *تراكبية المعنى* الذي يسم المركبات المعجمية المركبة يطغى على الواجهة.

ومقارنة، ولتسليط الضوء، يبتغى معاينة المقاربة الثانية من حيث أنها سلفا لا تطرح موضوع الصيغ الجامدة أو الوحدات المعجمية

---

[1] **Arsène Darmesteter** **عالم لغات ومستشرق فرنسي 1846- 1888. كان موجها لدراسات التلمود ولكنه غير توجهه إلى الدراسات اللغوية حتى عد من رجالاتها. من أعماله المشهودة في المجال *برسالة في تكون الكلمات المركبة في اللغة الفرنسية ومقارنة باللغات الرومانية واللاتينية* 1873، *وفي الابتكار الحالي للألفاظ وقواعده* 1877 ، *وحياة الكلمات من معانيها* 1887، وبالاشتراك مع آخرين *دروس النحو التاريخي للغة الفرنسية* 1891-1897،و *المعجم العام للغة الفرنسية من القرن السابع عشر حتى الآن*، 1895-1900.**

[2] **قضية اللوائح المغلقة لها أهمية كبرى في المعالجات الآلية والسؤال هو كمُّ هذه اللوائح المغلقة في أية لغة.**

المركبة، بل على العكس، إذ يتم في هذه المقاربة تحديد مجمعات الكلمات خلال التحليل ولا تمنح منزلة للمركب إلا في النهاية إذا اقتضى الحال لا غير. هذا التيار الثاني يولي أهمية لدراسة عملية الأداء المعجمي لنص أو متن بواسطة طرق إحصائية لتعداد الكلمات والبحث عن المتلازمات والهدف من ذلك دراسة أسلوبية أو خطابية للنص أو لكتاب أو كاتب وعموما لمتن معين.

والتياران البارزان في هذا التوجه وهما متأصلان من منابع مختلفة يشتركان في الاختيارات التالية:

1) أسبقية النص، ويتعلق الأمر بدراسة عملية *الأداء المعجمي* لنص أو متن أي أن التحليل الأسلوبي أو الخطابي هو في الصدارة؛

2) وتكون الكلمة في هذا التوجه هي وحدة التحليل الأساسية مستثنيا كل وحدة أخرى كالمقطع أو الجملة؛

3) ولا تشكل سلفا، حسب هذه الرؤية، مجمعات الكلمات وحدات معجمية في اللغة فدورها في النص هو الذي يحدد تجمعها وتكتلها؛

4) استعمال الطرق الإحصائية يمكّن حساب *المتلازمات* أو *المترابطات الانتظامية* أي البعد الفارق ما بين الكلمات بالمقارنة إلى كلمة محورية معتبرة. ويجدر بالذكر هنا أن استعمال التحليل الشعاعي في اللغة كانت مبكرة عند الإنجليز ويعود ذلك لـ 1960 .

وقد تتشعب التيارات والمواقف حسب اعتمادها أم لا على نظرية *لمعاني الكلمات* و حسب موقف التيار وتنظيره لاستعمال بطاريات التحليل الإحصائي في دراسة الكلمات وأيضا بحسب نوعية التطبيقات المقصودة.

أ) مدرسة لندن والمترابطات الانتظامية والمتلازمات

جاء تحليلها كوليد لتحليل المترابطات الانتظامية في *النظرية السياقية لمدرسة لندن*[1] واتخذت من مقالة جون ريبرت فورث[2] المنشورة عام1957 في أعماله منطلقا لتحليلاتها، وكان هذا النص قد ميز مستوى تحليل لساني خاص أطلق عليه التعبير *بالمترابطات الانتظامية* :

التعبير عن معنى بالمترابطات الانتظامية هو عملية تجريد على المستوى المقطعي وليس مرتبطا مباشرة بمقاربة تصورية لمعنى الكلمات. مثلا كلمة "ليلة" ترتبط وتنتظم مع كلمات "سوداء" و"ظلماء" و"بهيم" و"بهيم" ذاتها تنتظم مع "ليل"[3].

ما يسميه فورث بالترابط الانتظامي يعني، في الواقع، عبارات وصيغا وتراكيب ومقاطع مشكلة في النص انطلاقا من كلمة محورية أو مفتاح. ولم يتطرق فورث في كلامه لسمة التكرار ولا لتحديدها بالطرق الإحصائية وذلك عكس ما تؤكده بعض دراسات *اللسانيات المتنية* الحالية من كون فورث كان سباقا للميدان. ونحيل هنا لمتابعة مراحل تاريخ مدرسة لندن لكتاب ستوب[4] ؛ وفي الحقيقة إن الأبحاث حول الترابط الانتظامي عند فورث تدخل في إطار الدراسات الأسلوبية. وقد استغل طلاب فورث فكرة التعبير عن المعنى بالترابط المنتظم في الطرق

---

1 يطلق اسم مدرسة لندن على ما خلقه فورث من تأثير وتيار سمي بتيار فور تابعه طلابه كمكايل هاليداي وميتشيل وبالمر وغيرهم وعرف من أنصار هذه المدرسة في العالم العربي كل من إبراهيم أنيس وحسن تمام.

2 John Rupert Firth « Modes of meaning », 1951 -1959، 1960 -1980 لساني إنجليزي له توجه وتصور مشهور. قال بأن معنى الكلمة يعرف بسياق ترابطها، سياق الموقف. وأن اللغة أنظمة متعددة يحددها الموقف.

3 النص نقل لفكرة الباحث وليس حرفيا.

4 Stubbs *Corpus Linguistics*, 1993

الآلية بداية سنوات الستينات وخاصة طلابه ميكايل هاليداي[1] في الترجمة الآلية عام 1956 وجون سينكلير عام 1966 وجونس وسينكلير[2] عام 1974 في المعجمية. وكان لموضوع معالجة المعجم آليا تأثير كبير على المدرسة البريطانية ونلاحظ أن بعضهم طور أبحاثا في المعالجة الآلية تستمد أرضيتها من تحاليل الارتباطات الانتظامية.

ب) المدرسة الفرنسية

في السنوات التي تهمنا ظهرت الدراسات الإحصائية مع أبحاث ماريو روك[3] وجرده العام للغة الفرنسية في سنة 1936 ويمكن أن نذكر للتأريخ خلق مختبر المعجميات في جامعة بيزونسون[4] الفرنسية في عام 1959 وإصدار الدوريتين : دفاتر المعجمية ودراسات اللسانيات التطبيقية [5] وكذلك أبحاث شارل ميلير[6] الذي كان يهدف إلى تشجيع المعجميات الإحصائية كجزء كائن بذاته متضمن في اللسانيات. وحققت الدراسات الإحصائية زخما نوعيا مع خلق مختبر المعجميات السياسية في المدرسة الوطنية العليا الفرنسية بسان كلود في أواخر الستينات.

ومكانة العناصر المتلازمة في أبحاث الحسابات المعجمية لم تطرح إشكالا إلا في مراحل تالية للأبحاث، فالأبحاث الأولى التي استعملت الطرق الإحصائية الاحتمالية مزجت بين نوعين من الارتباطات المختلفة:

- أزواج موظفة داخل وحدات معجمية مركبة؛

---

[1] **Michael Halliday**
[2] **John Sinclair & Jones**
[3] **Mario Roques**
[4] **Besançon**
[5] ***Les cahiers de lexicologie & Les études de linguistique appliquée***
[6] **Charles Muller**

- و أزواج موجودة في نفس الجمل ولكن على تباعد متغير.

مثلا لا مجال للحكم على عبارة الشركات متعددة الجنسيات في نص يحوي 12 تكرارا لكلمة شركة خمس منها فقط متبوعة بمتعددة الجنسيات؛ ولذلك ابتكر الباحثون طريقة أسموها جرد المقاطع المكررة والتي تحدد الأشكال المتشابهة؛ ولكن مع ذلك من بين هذه الأشكال نجد في نفس الوقت المتلازمات "donner le coup de grâce"، أطلق رصاصة الرحمة، ومقاطع اسمية حرة " la colère du père Duchesne"، غضب الأب ديشين، وأيضا متواليات بلا استقلالية تركيب ووظيفة وصعبة التأويل مثل "n'y a pas de "، لا يوجد ...؛ وهي أشكال وتراكيب كما نرى متنوعة المنزلة[1] فمنها ما هو مترابط جامد ومنها ما هو حر التركيب أو مفتوح؛ فأي وضع يمكن منحه لهذه الأشكال؟ وترك تأويل كل هذه الإشكالات في هذه المدرسة للتحاليل التي تهتم بالمتن. ونسجل بأنه في فرنسا بدأ التحليل الآلي للغة بتأليل المعجم والذي كان يدخل في صلب اهتمامات عديد من لساني تلك الفترة . أي دراسة المعجم الفرنسي والأسلوبية . وهذا عكس الوضع الإنجليزي حيث كانت نظرية معنى الكلمات منبع أوائل الأبحاث في الترجمة الآلية والدراسات المعجماتية. وهما رؤيتان ستندمجان مع تجدد الدراسات الإحصائية سنوات 1990 تحت إغراء العودة لمقولات فورث. وفعليا لم تكن قطيعة مطلقة بين الرؤيتين طوال تاريخهما، فمثلا طلاب فورث هاليداي وسينكلير شاركا في أبحاث مركز بيزونسون. وتجدر الإشارة إلى أن ارتباط المقاربتين بالبنيوية لم تكن بنفس الدرجة ويتجلى ذلك في مسألة التمييز بين التحليل البنيوي والتحليل الإحصائي.

---

[1] **احتفظنا بأمثلة فرنسية لأننا بصدد الحديث عن التجربة الفرنسية، ولنا في العربية أمثلة مشابهة.**

ج) التجربة الفرنسية والمعالجة الآلية.

مشاريع الترجمة الآلية التي رأت النور في فرنسا في الستينات شجعت اللسانيين على إعادة الاعتبار للكلمة كوحدة تحليل، علما بأنها كانت منبوذة بل عنصرا مشبوها في الرؤية البنيوية. فجورج مونان[1] مثلا يذكر بأن كل تعريف للصيغ المسبوكة كواقع لغوي تستمد كلها ركيزتها من سمات سلبية بحتة وهي من متبقيات التحاليل الصورية. وبالتوازي مع اعتماد الكلمة المكتوبة كوحدة أساسية للتحليل النصوص، . والكلمة المكتوبة هي سلسلة أحرف بين فارقين يطرح سؤال حول منزلة مجمعات الكلمات هل هي وحدة متفردة أم متعددة العناصر؟ إن تراث الدراسات المعجماتية وتأثير الاهتمام الذي رافق نظرية شانون وويفر[2] الرياضية للتواصل من جهة، و خصوصيات اللغة الفرنسية كلغة ذات مستوى تكتل إدغامي منخفض وتتطلب استعمال عدة كلمات لتشكيل كلمات مركبة من جهة أخرى، هي وإلى حد ما من أسباب اعتبار الكلمة المكتوبة كوحدة. ويمكن اعتبار أبحاث برنارد بوتيي[3] بالخصوص لأنه واحدا من اللسانيين الذي اهتموا بالمعالجة الآلية ولأنه فتح الطريق لعديد من الاجتهادات المتعلقة بمجمعات الكلمات من نوع ما تطرق إليه إيميل بن فينست[4] و أندري مارتيني[5]. إن ضرورات ترجمة نص آليا

---

[1] Georges Mounin 1910-1993 ، اسم شهرة للساني فرنسي تناول بالبحث عديدا من حقول اللسانيات ومن منشوراته المعروفة *القضايا النظرية في الترجمة* 1965،ومع آخرين *معجم اللسانيات* 1974.

[2] Shannon & Weaver كلود شانون 1916- 2001 باحث في مختبرات بيل مهندس كهربائي؛ وارين ويفر 1896- 1978 رياضي أمريكي. عرفا بنظريتهم الرياضية للتواصل.

[3] Bernard Pottier 1924- لساني فرنسي له مساهمات مرموقة ومن منشوراته *التمثيل الذهني والتصنيف اللساني* 2000، و *علم الدلالة العام* 1992.

[4] Émile Benveniste 1902- 1976، لساني فرنسي اشتهر بأبحاثه في النحو المقارن للغات الهند أوروبية واللسانيات العامة. ومن كتب المعروفة قضايا في *اللسانيات العامة*.

[5] André Martinet 1908- 1999، لساني فرنسي. تأثر بمدرسة حلقة براج اللسانية. وشيد المدرسة الوظيفية في التركيب، وصاحب مبدأ اقتصاد التغيرات اللسانية ومكتشف

دفعت ببوتيي لتحديد وحدات تكون كلا من وحدات تقطيع للنص وللغة وللمعنى والترجمة يعني يجب أن يتماثل الشكل المكتوب مع الوحدة التركيبية ومع الوحدة الدلالية. وللجواب على هذه الشروط المتعددة طرح بوتيي ترتيبا للكلمات ومجمعات الكلمات في إطار وحدات معجمية وهي وحدات وظيفية محددة في اللغة. فزيادة على الوحدات المعجمية البسيطة "كفرس" والمركبة كقولك في الفرنسية "sous-groupe"، مجموعة جزئية[1] يقول بوتيي بوجود وحدات معجمية معقدة مثل "توت الأرض"[2] والتي ستطرح مشكلة تعرف للآلة التي يجب عليها أن تميزها عن مجمعات الكلمات الحرة أو مقاطع. وقد اختار بوتيي عموما شروط تحديد تركيبية وتسِم تلك الشروط انغلاق الوحدة عن أي شكل من الإلحاق، كشروط اللا تفارق واللا تخصيص واللا تكميم، إضافة إلى شرط دلالي منطقي وهو اللا إسناد ؛ وقد اهتم بوتيي أيضا بنمط تشكل الوحدات المعجمية المعقدة وذكر بنيتها الداخلية مصنَّفة مثل: مسميات أوصفات أو أفعال حيث نرى بوادر معالجة تركيبية؛ وأدخل تصنيفية تشابه فكرة *الرأس* في *النحو المقطعي* فقد صنف مثلا "الدار الفانية" كمسمى لوجود صنف أعلى هرميا : مسمى + صفة = مسمى[3]. غير أن شرطا من الشروط التي يذكرها بوتيي هو دلالي :

=التمفصل المزدوج للسان البشري. له العديد من المنشورات منها *عناصر اللسانيات العامة* 1955، *والتركيب العام* 1985، *وظيفة وحركية اللغات* 1989.

[1] لا تمتلك اللغة العربية هذه الخاصية وهي الخط المدمج ففي اللغة الفرنسية يعني الخط المدمج بين كلمتين أنهما أصبحتا كتلة وكلمة واحدة وهي مقدمة لدمج الكلمتين كليا كقولك تأبط شرا وتأبطشرا.

[2] يتحدث مارتيني عن كلمات في اللغة الفرنسية من شاكلة pomme de terre، بطاطس و boite aux lettres ، صندوق البريد وهي وحدات معقدة وتراكيب كانت تطرح مشكلة للمعالجة الآلية. فعلا يمكن للآلة تحديد "توت" و "الأرض" ولكن "توت الأرض" شيء آخر.

[3] إذا كنا بحكم ثقافتنا اللغوية والاجتماعية نعرف ما تعنيه الدار الفانية وهي الحياة الدنيا وهي عكس الحياة الآخرة، فما سيكون فهمنا لمقطع الدنيا الخالدة أو الأمة الخالدة وقد صبتا في نفس القالب تركيبيا.

وهو هوية مجمع الكلمات وإحالته على مسميات ثابتة. وربما بتأثر ببار هيليل أدخل بوتيي شرطا آخر مرتبط بهدف الترجمة مصدره المقارنة ما بين اللغات : فوحدة معجمية معقدة قد تترجم بوحدة معجمية بسيطة في لغة أخرى.

ويزيد بوتيي شرطا إحصائيا مرتبطا بضرورة تصنيف وتدوين . تخزين[1] العبارات والصيغ المسبوكة في لوائح ومعاجم كشرط مسبق لأية معالجة آلية وقد طرح فكرة تصنيفها حسب مستوى معجميتها المحدد بالطرق الإحصائية؛ مثلا في اللغة العربية "ثالث" و"أثافي" ولنأخذ تطبيقا جمل "ثالثة الأثافي" و"ثالث معجزة".

نعرف جيدا بأن عددا كبيرا من الوحدات قد تلي كلمة " ثالث"، ولكن احتمال أن تسبق كلمةَ معجزة كلمةُ ثالث ضعيفة حسابيا على العكس من ذلك توجد قلة من الكلمات المتداولة التي قد تسبق كلمة" الأثافي "، واحتمالات أن تكون "ثالث" سابقة لكلمة أثافي عالية جدا؛ إذاك نقول بأن "ثالث الأثافي" لها مستوى عال من القابلية للمعجمة[2]. ولكن بوتيي لا يحدد مما يتشكل مستوى القابلية للمعجمة هذا، وقد نفهم من خلال كلامه بأن هذه القابلية مرتبطة بعدد التكرارات وأن الأخيرة تستنبط بمعاينة المتن : يقول بوتيي[3] ما مفاده لنجاح هذه المهمة يجب القيام بجرد آلي لاستخراج العلاقات والارتباطات المعجمية في الخطاب والتي توافق معجمية في اللغة؛ وقد يكون لفكرة الجرد هذه علاقة بما قاله أصحاب مؤلف :*الفرنسية الأساسية*[4] وكانوا يهدفون من

---

[1] لا يتعدى الأمر جمعا وتدوينا لمعطيات وتستعمل كلمة تخزين في المعالجات الآلية.

[2] وما (ضيع الصيف) إلا .... (اللبن). وما أكثر الأمثلة التي تفسر حساب الاحتمالات هذا.

[3] بوتيي صفحة 205 ، 1960

[4] **Gougenheim, Rivenc , Michea & Sauvageot, *Le Français élémentaire*, 1954.**

ذلك تقديم تعليم اللغة الفرنسية كلغة أجنبية وهي طريقة مبنية على عدد تكرار الاستعمالات *والمتاح المعجمي* وقد كونوا مؤلفهم المذكور باستعمال متن شفوي من مختلف المستويات والأصناف وبلغت قائمة الكلمات في مؤلفهم 1350 كلمة مكررة أو متاحة، وقد جمعت القائمة بعد فرز 163 حوارا مع 75 ناطق؛ ويعد عملهم رياديا في المجال. وما يهمنا هنا، وهو ما يطرح إشكالا، هو مشكل النقل والانتقال من المتن إلى اللغة للخصائص الإحصائية للمعجم دون إبراز شرعية وعلل الانتقال، وهذا الانتقال كان محط جدل في الستينات ولازال عند أصحاب *اللسانيات المتنية*، فمثلا شارل ميلير يميز بين الكلمات والمعجم ويخصص الأولى للغة والثانية للخطاب وعلى هذا الدرب سار توراتيي. وتبرز أبحاث بوتيي التداخل فيما بين التحليل البنيوي والتحليل الإحصائي والتي تصدرت أبحاث التأليل الفرنسية في سنوات الستينات وتسلط الضوء أيضا على الصعوبات المرتبطة بتداول المتون وتعميم نتائجها على اللغة.

د) بار هيليل والترجمة الآلية

كان بار هيليل أو باحث وظِّف للعمل حصرا على مشروع للترجمة الآلية بالمعهد التقاني بمساشوشيت[1] سنة 1951 ويعد بار هيليل أول والوحيد الذي اهتم بالمعالجة الآلية للصيغ المسبوكة في الولايات المتحدة وذلك منذ 1951 معتبرا بأن تصنيف تلك الصيغ وتدوينها في معاجم هو شرط مسبق لأية عملية في الترجمة الآلية (وقد أشار جريماس[2] إلى توجه بار هيليل مبكرا في فرنسا في مقال له عام 1960 خصه لصعوبة تصنيف الصيغ المسبوكة). وحول هذه النقطة تميز بار

---

1 MIT.
2 Greimas

هيليل عن البنيويين الأمريكيين الذين تجاهلوا كليا الوحدات المعجمية المعقدة، وهذا التجاهل ليس مستغربا إذا وضعنا في ذهننا بأن *الصيغم* هو الوحدة الصغرى وليس الكلمة في النحو التوليدي وأن وحدة التحليل الأساسية هي الجملة وبالفعل لم يمنح النحو التوليدي أهمية للمعجم إلا بعد أن أدخل أرنوف[1] في سنة 1956 *مكونا صرفيا اشتقاقيا*. واختلف بار هيليل عن البنيويين أيضا بتبرئه من استعمال *طرق الاستكشاف* لتحديد الأصناف اللسانية، فهو يرفض فعليا 'اعتبار الصيغ المسبوكة كأصناف تامة وطرح إمكانية تعريفها تعريفا متغيرا بحسب الأهداف المبتغاة؛ وحيث أنه يتموقع في زاوية تستهدف الترجمة الآلية فقد شدد على عدم قابلية الصيغ المسبوكة للترجمة نظرا' لخصائصها غير التركيبية مثلا صيغة rouge gorge الفرنسية وجوز الهند وإلى حد ما قميص عثمان العربيتين يمكن ترجمتها إلى مختلف اللغات ولكن أيا من مكونات الصيغة منفردة تسمح لنا بترجمة مقبولة، غير أنه يمكن استعمال وتوظيف حرفي وعادي للصيغ[2] وهو ما يؤكد عدم كون الصيغ المسبوكة صنفا لسانيا قائما بذاته ولكنها من دائرة الخطاب أيضا. علما بأن بار هيليل لم يطرح مشكلة المَعجمة لا في اللغة ولا في الخطاب فهو ينطلق من الوحدات المعجمية وهي الصيغ المسبوكة ويختار التدوين والتخزين مسبقا.

ه) مدرسة هاريس

وإذا كان البنيويون الأمريكيون لم يولوا أهمية للوحدات المعجمية المركبة ففي فرنسا فقد وظِّف *التحليل التوزيعي* في الأبحاث حول تكتل

---

[1] **Aronoff**

[2] **وما يمنعا من الحديث عن قميص عثمان وهو حي يرزق بيننا غير الذي نعرف تاريخيا وغير استعمال شاع وتعمم مثلا، زد لذلك أن قميص عثمان واقعة تاريخية والمؤرخ يستعمل قميص عثمان الذي نقل للأمصار استعمالا عاديا.**

الكلمات اعتمادا على أبحاث هاريس غير أن هذا الاعتماد ولد تيارين مختلفين تيار جون ديبوا وتيار موريس جروس[1]. فجون ديبوا على عكس بوتيي لا يهتم بأنماط التشكيل ولا قابلية تكتلات الكلمات للمعجمة؛ فهو يقاربها كوحدات دلالية، وبرأيه فإن تحديد المجموعات الاسمية المعقدة ليس لها إلا وظيفة خطابية في متن معين. وفي دراسة للمعجم السياسي والاجتماعي الفرنسي بين سنوات 1869 و 1872، نشرت عام 1960، يفحص ديبوا الوحدات المعقدة مثل : تحرر العمال، "émancipation des travailleurs[2]". فالدور الذي تلعبه هذه المجمعات في المعجم الثوري تفقدها خصوصيتها التركيبية لصالح سمة دالة جديدة والتي تضاف أو تعوض السمات الأخرى "كإيجاد نظام اقتصادي جديد مفيد للعمال" ويستعمل مصطلح *التحييد* وهو مصطلح *صواتي* لوصف عملية تحول الوحدات المعجمية المركبة إلى وحدات دلالية إذا لم نقل خطابية. وانطلاقا من الأدبيات الاجتماعية لفترة الدراسة يستخرج ويكون الباحث سلسلة من الوحدات الدلالية انطلاقا من *المجموعة المتعادلة* :

تحرير + (العمال/ الجماهير/ الشغيلة/ الطبقة الكادحة/ البروليتاريا).

في هذه العبارات يلعب العنصر الثاني دور المتغير الذي يمكن تبديله دون الإخلال بالمعنى الدلالي لكل الوحدة والعناصر التي هي في موقع مهم هي التي تلعب دور المتغير وحيث أنها متبادلة فهي محيَّدة. ويؤكد ديبوا على الخصائص الخطابية للمجموعات المتبادلة. وفعلا ففي نفس الفترة وفي سياقات أخرى كان مستعملو اللغة لا يوظفون بنفس

---

[1] **Jean Dubois & Maurice Gross**

[2] **احتفظنا بمثال ديبوا الفرنسي لأننا إزاء نفس المعطى في العربية.**

التساوي كلمات الطبقة العاملة والعمال و الجماهير ولم يكن ممكنا بتاتا تبديل كلمة البروليتاريا بكلمة العمال. ونسجل الخصوصية الحركية للوحدات المعقدة التي يعرضها ديبوا. ففي الخطاب تأخذ الوحدات المعقدة دلالة خاصة مقارنة المقطع الحر المقابل إما بالزيادة أو بالتعويض والتحييد. ونرى هنا بوادر مدرسة تحليل الخطاب الفرنسية التي اعتمدت على التحليل الهارسي والتي سترى التأليل في أبحاث ميسال بيشو[1] عام 1969.

وتندرج في خضم تيار البنيوية الأمريكية وبتأثير مباشر من هاريس[2] أبحاث موريس جروس[3] ومختبر معالجة الوثائق اللسانية على الجمل الجامدة (الصيغ الجامدة) والتي ستجد طريقها للتأليل فعليا في سنوات الثمانينات. وقد اختبار جروس نظريا منهجا يعتمد على التجريبية فهو يرى بأن دراسة منتظمة لكل العناصر تمكن من اكتشاف الخصائص العامة المحتملة لذلك هدف جروس لجرد شامل للصيغ الجامدة تجسد في فكرة خلق معجم نحوي كأساس لمعجم رقمي. وفيما يتعلق بهذا الجرد للصيغ الجامدة يبرز تأثير الهاريسية في عدة جوانب :

1. الاهتمام بالتصنيف وتكوين مجموعات تعادلٍ دلالية بطرق تركيبية؛
2. استعمال عبارة صيغ جامدة يعكس الأسبقية المولاة للجملة البسيطة كوحدة وليس للكلمة وذاك يعبر عن اختيارات هاريسية؛
3. الأهمية المولاة للتركيب، فخصائص تعريف الجمل والصيغ الجامدة هي حصرا تركيبية، و *اختبارات* إمكانيات تطبيق *التحويلات*

---

1 Michel Pêcheux

2 Zellig Sabbetai Harris 1909 – 1992. كان خالق أول شعبة للسانيات بالولايات المتحدة عام 1946؛ وعرف بنظريته في النحو التحويلي وكان أول من صورن التحليل التوزيعي. وشجع على ريضنة وصورنة التحاليل اللسانية بلغة سهلة.

3 Maurice Gross 1934 - 2001 لساني فرنسي تأثر بهاريس وله نظرية المعجم - النحو.

تخص أساسا مجال الأفعال؛ ويميز جروس بين ثلاثة أنواع جمل أساسية حسب تشكلها على ركيزة هي إما أفعال عادية أو أفعال تدخل في الصيغ الجامدة أو أفعال سند.

ويتقاسم أيضا جروس فكرة بار هيليل، والتي تبناها جريماس، والتي تقول بأن الصيغ الجامدة لا يمكن أن تشكل صنفا لسانيا خاصا قائما بذاته، ففي تصور موريس جروس توجد استمرارية غير متقطعة بين الأشكال الحرة المتغيرة والأشكال الجامدة التي لا تتغير، وتمتاز الجمل الجامدة بترابط تركيبي أقوى مما في الجمل الحرة ولكنه ترابط يمكن دائما تغييره[1]. ويدافع جروس عن تخزين كصيغ جامدة كل الجمل الني تتضمن ما يسميه الخصوصيات النحوية وكذا الصيغ الجامدة حقا. ومن المفارقات أننا نحصل بذلك على عدد من الجمل الجامدة أكبر من عدد الجمل الحرة ولذلك يجب تدوينها في لوائح معجمية قبل أية معالجة آلية. وتجدر الإشارة أن جروس لا يتناول الصيغ المعجمة سابقا والغير متأثرة بتغيرات الخطاب. وإذا كان ممكنا دائما تحليل صيغة جامدة إلى مقاطع حرة فالعكس ليس صحيحا مطلقا، فمقطع حر إذا لم يخزن سلفا لا يمكن أن يعتبر صيغة جامدة.

و) المقاربة الدلالية لإيجور ملتشوك

جاء اهتمام ميلتشوك[2] بالوحدات المعجمية المركبة كمتابعة لأبحاثه في روسيا على الترجمة الآلية في الخمسينات وكانت أبرز محطات أبحاثه نظريته في *المعنى والنص ومعجمه المفسّر والمؤلِف للغة الفرنسية المعاصرة*، وتختلف مقاربة ميلتشوك جذريا عن المقاربات

---

[1] ضيعت الصيف اللبن واللبن أضاعه الصيف وضاع اللبن مع الصيف وضاع اللبن وأضيع اللبن. وخبأ قرشك الأبيض وهي كلها تعبير عن تغيير تركيبي لوحدات معجمية.

[2] Igor A. Mel'cuk ولد بروسيا 1932 واشتغل بالمعهد علوم اللغويات الروسي على نظرية المعنى والنص قبل أن يهاجر عام 1976 وهو حاليا أستاذ الترجمة بجامعة مونريال ومن مؤلفاته زيادة على المعجم المذكور دروس التصريف العام.

السابقة بالتأكيد على عملية تشكل وإنتاج الوحدات المعجمية المركبة عوض الاهتمام بالتعرف عليها وعلى خصائصها الدلالية الداخلية، ومنذ سنوات السبعينات انصب اهتمام ميلتشوك أكثر على الاهتمام بالتركيب اللغوي في اللغة هدفِ الترجمة الآلية منه بالاهتمام بتحليل اللغة المصدرية، ويفترض مع طرح أسسٍ لغةً وسطية مبنية على التمثيل الدلالي المشترك بين اللغات؛ وكذلك الأمر بالنسبة للصيغ الجامدة إذ يولي ميلتشوك أهمية لإنتاجها بالبحث عن السمات اللسانية المعنوية والتي تعلل اختيار الناطقين لها. ويميل ميلتشوك في دراسته للصيغ الجامدة للارتكاز على المقطع وليس على الكلمة كوحدة أساسٍ في التحليل فهو يرى بأن شيوع المقاطع الجامدة في اللغة أكثر من شيوع المقاطع الحرة وأن الناطقين بلغة يتحدثون مستعملين المقاطع الجامدة أكثر من استعمالهم للكلمات وهو موقف يشبه مواقف جروس الذي يشترك مع ميلتشوك في فكرة إجبارية الناطقين على حفظ الصيغ الجامدة وأن هذا الحفظ مرتبط بالجانب الإدراكي منه بالجانب اللساني. ويعتبران أن كل معالجة آلية هي مشروطة بتكوين معجم شامل إذا أمكن مسبق للمقاطع الجامدة غير أنهما يختلفان في مواقف أخرى فميلتشوك لا يفترض تتابعا خطيا بين المقاطع الحرة والمقاطع الجامدة وزيادة على ذلك فهو يرتب المقاطع الجامدة دلاليا وليس تركيبيا ويميز بين أربع أنواع من المقاطع حسب مستوى تراكب دلالة مختلف المكونات:

1. المستوى الذريعي "يستهلك قبل تاريخ" وهي مقاطع ذات معنى واضح وتركبي ولكنها حسب اللغات تكون على أشكال توافقية ليست قابلة للتبديل؛

2. مستوى المقاطع الجامدة التامة "أكذب من مسيلمة" غير تراكبية ولا قابلة للترجمة؛

3. مستوى شبه المقاطع الجامدة أو المرتبطات الانتظامية "عمل خيرا" أو العبارات المعجمية الوظيفية فالمعنى تركبي وتكون الكلمة المفتاح متبوعة بعنصر مسند؛

4. مستوى شبه المقاطع الجامدة وشبه الصيغ الجامدة "يوم الحساب" موسومة بتركب المعنى وإسنادية العنصر المضاف.

من خلال ما سبق عاليه نرى بأن المعالجة الآلية للوحدات المعجمية المركبة مكنت من إبراز أن التعارض المتفق عليه حاليا بين التحليل البنيوي والتحليل الإحصائي ليس تعارضا موفّقا وفعالا فقد لاحظنا مثلا بأن بوتيي يستعمل التحليل الإحصائي لتحديد مستوى المعجمية دون توضيح الغموض الحاصل بين مستويات المتن واللغة وأن ديبوا حين يبرز المتلازمات في متن دون سند إحصائي ولكن مرتكزا فحسب على منطلقات خطابية بحتة. ولاحظنا أن مواصفات دلالية بحتة يمكن أن توظف لترتيب وتخزين الوحدات المعجمية المركبة. ومن الجلي أيضا بان المعارضة بين تحليل بنيوي وتحليل إحصائي لا تغطي كامل التعارض بين اللغة والخطاب أو بين اللغة والمتن. وبعض الأبحاث الجديدة في فرنسا التي ميزت بين نمط التشكيل والمعجمة واستهدفت في أبحاثها تحديد الوحدات المعجمية المركبة بواسطة *محلل تركيبي* وإمكان اعتبارها جامدة بالنظر لمتن معين. وأخيرا يذكر إيفيراريت[1] وزملاؤه أن الأبحاث المتأخرة تستكشف *مسار التعرف* على *إنتاج الصيغ* الجامدة انطلاقا من فرضيات إدراكية أو لغويات نفسية.

ومن عاليه نأمل أن يكون المتتبع قد فهم جدوى هذا النقاش حول الوحدات المعجمية المركبة وفطن إلى أهميتها في مجال تأليل

---

[1] Everaert

اللغات ومنها اللغة العربية علما بأننا لم نتطرق لمناقشة المرتكزات النظرية لوجود أو لا للوحدات الجامدة وهو نقاش لا يهم في عمقه الباحث في مجال التأليف فما يهم المؤلل هو فعالية النتيجة قبل كل شيء.

المراجـــــــــع :

ARONOFF Mark, *Word formation in Generative Grammar*, "Linguistic Inquiry Monographs I", Cambridge, MIT Press, 1976.

BAR-HILLEL Yehoshua, "Idioms", *Machine Translation of Languages, 14 essays*, pp: 183-193, William N. Lock & Andrew D. Booth éd., New York, MIT et John Wiley, 1955.

DAVID Sophie, « Remarques à propos du monde de construction des unités de forme NN », *TAL*, n° 34-2, pp : 59-74, 1993.

DUBOIS Jean, « Les notions d'unité sémantique complexe et de neutralisation dans le lexique », *Cahiers de lexicologie*, n° 2, pp : 62-66, 1960.

EVERARAET M., VAN DER LINDEN E. -J., SCHENK A. & SCHREUDER R. éd., *Idioms: Structural and Psychological Perspectives*, Hillsdale, Lawrence Erlbaum Associates, 1995.

FIRTH John Rupert, « modes of meaning », *Papers in linguistics (1934-1951),* pp: 190-215, Oxford University Press, 1957[1951].

FRANÇOIS Jacques et MEJRI Salah éd., *Composition syntaxique et figement lexical*, Coll., « Bibliothèque de la Syntaxe & Sémantique », Presses universitaires de Caen, 2006.

GOUGENHEIM G. MICHEA R., RIVENC P & SAUVAGEOT A., *L'élaboration du français élémentaire*, Paris, Didier, 1954.

GREIMAS A.-J., « Idiotismes, proverbes dictons », *Cahiers de lexicologie*, n° 2, pp : 41-61, 1960.

GROSS Maurice, « Les limites de la phrase figée », *Langages*, n° 90, pp : 7-22, 1988.

GROSS Maurice, « Le programme d'extension des lexiques électroniques », *Langue française*, n° 87, pp : 123-127, 1990.

GROSS Maurice, « Les phrases figées en français », *L'information grammaticale*, n° 59, pp : 123-127, 1993.

HALLIDAY Michael A.K., « Linguistic basis of the thesaurus-type mechanical dictionary and application to English-preposition classification" *Mechanical translation* n° 3-2, pp: 36-37.

LAFON P. & SALEM A., "conceptions du mot et débuts de la traduction automatique", *Mots*, n° 6, pp: 161-177, 1983.

LÉON J., « Conceptions du mot et débuts de la traduction automatique », *Histoire Épistémologie Langage*, n° 23, pp : 81-106, 2001.

LÉON J., « *Lexies, synapsies, synthèmes* : le renouveau des études lexicales en France au début des années 1960 », *History of Linguistic in Text and Concepts*, pp : 405-418, Gerda HASSLER éd., Nodus Publication, 2004.

MEJRI Salah « Figement et dénomination », *Meta* n° 4, vol. 45, pp : 609 – 621, 2000.

MEL'CUK Igor A., ARBATCHEWIKY-JUMARY N., ELNITSKY L., IODANSKAJA L., & LESSARD A., *Dictionnaire explicative et combinatoire du français contemporain*, Montréal, Les Presses de l'université de Montréal, 1984.

MOUNIN Georges, *La machine à traduire. Histoire des problèmes linguistiques*, La Haye, Mouton, 1964.

MULLER Charles, *Initiation à la linguistique statistique*, Paris, Larousse, 1968.

PÊCHEUX Michel, *Analyse automatique du discours*, paris, Dunod, 1969.

POTTIER Bernard, « Les travaux lexicologiques préparatoires à la traduction automatique », *Cahiers de lexicologie*, vol. 2, pp : 200-206, 1960.

SINCLAIR J. & JONES S., "English lexical collocations, A study in computational Linguistics", *Cahiers de lexicologie*, n° 24, pp: 15-61, 1974.

STUBBS Michael, « British traductions in text analysis-from Firth to Sinclair", Text and technology. In Honour of John Benjamins, pp: 1-36, Francis & Elena TONINI-BONELLI éd., Amesterdam, John Benjamins, 1993.

TOURNIER Maaurice, "D'où viennent les fréquences de vocabulaires?", *Mots*, n° 1, pp : 189-209, 1980.

WILKS Yorik A., SLATOR B. & GUTHRIE L., *Electric Words: Dictionnaries, Computers and Meanings*, Cambridge, MA, Londres, MIT Press, 1996.